yang_seonhyeHW8

```r
library(data.table)
data <- fread("http://users.stat.ufl.edu/~winner/data/hotel_energy.csv")
hotel_energy <- data.table("hotel_energy.csv", header = TRUE)
```

Question 1

Since we want to test the hypothesis where $\beta_{area} = \beta_{age} = \beta_{numrooms} = \beta_{occrate} = 0$

```r
attach(hotel_energy, warn.conflicts = F)
model <- lm(enrgcons~area+age+numrooms+occrate, data=data)
summary(model)
```

```
##
## Call:
## lm(formula = enrgcons ~ area + age + numrooms + occrate, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3627416 -1462717   622058  1052335  2490428
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.692e+06  2.717e+06  -1.727   0.1061
## area         2.071e+02  3.765e+01   5.500 7.82e-05 ***
## age          1.562e+04  1.331e+05   0.117   0.9082
## numrooms    -7.226e+03  6.026e+03  -1.199   0.2504
## occrate      7.367e+06  3.867e+06   1.905   0.0776 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1949000 on 14 degrees of freedom
## Multiple R-squared:  0.8391, Adjusted R-squared:  0.7931
## F-statistic: 18.25 on 4 and 14 DF,  p-value: 1.92e-05
```

```r
reduced_model <- lm(enrgcons~1, data=data)
anova(reduced_model, model)
```

```
## Analysis of Variance Table
##
## Model 1: enrgcons ~ 1
## Model 2: enrgcons ~ area + age + numrooms + occrate
##   Res.Df        RSS Df  Sum of Sq      F   Pr(>F)
## 1     18 3.3035e+14
## 2     14 5.3155e+13  4 2.7719e+14 18.252 1.92e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Looking at the result, there is evidence to conclude that at least of the variables ($\beta_{area}, \beta_{age}, \beta_{numrooms}$, and $\beta_{occrate}$) is not equal to zero.

Question 2

```r
model <- lm(enrgcons~area+age+numrooms+occrate, data=data)
reduced_model1 <- lm(enrgcons~numrooms+occrate, data=data)
anova(reduced_model1, model)
```

```
## Analysis of Variance Table
##
## Model 1: enrgcons ~ numrooms + occrate
## Model 2: enrgcons ~ area + age + numrooms + occrate
##   Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
## 1     16 1.6862e+14
## 2     14 5.3155e+13  2 1.1546e+14 15.205 0.0003094 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Looking at the result, there is evidence that at least one of β_{area} and β_{age} does not equal to zero.

Question 3

Since X_1 and X_2 are independent, we can find the pdf of X_1 and X_2 and multiple them.

we have the distribution of chi-square distribution:

$$\chi^2 = \frac{x^{n/2-1}e^{-x/2}}{2^{n/2}\Gamma(n/2)} \qquad n = df$$

for pdf of f(x):

$$f(x_1) = \frac{x_1^{d_1/2-1}e^{-x_1/2}}{2^{d_1/2}\Gamma(d_1/2)}$$

for pdf of f(y):

$$f(x_2) = \frac{x_2^{d_2/2-1}e^{-x_2/2}}{2^{d_2/2}\Gamma(d_2/2)}$$

Since $X_1 = \chi^2(d_1)$ and $X_2 = \chi^2(d_2)$, we just take the df and plug it into the distribution to get the pdf.

Joint pdf is

$$f(x_1) \cdot f(x_2) = \frac{x_1^{d_1/2-1}e^{-x_2/2}}{2^{d_1/2}\Gamma(d_1/2)} \cdot \frac{x_2^{d_2/2-1}e^{-x_2/2}}{2^{d_2/2}\Gamma(d_2/2)}$$

$$f_{X_1X_2}(x_1, x_2) = \frac{e^{-\frac{(x_1+x_2)}{2}}x_1^{\frac{d_1}{2}-1}x_2^{\frac{d_2}{2}-1}}{2^{\frac{d_1+d_2}{2}}\Gamma(\frac{d_1}{2})\Gamma(\frac{d_2}{2})}$$

Question 4

Joint PDF of x_1 and x_2 is given by:

$$f_{x_1x_2}(x, y) = \frac{e^{-\frac{(x+y)}{2}}x^{\frac{d_1}{2}-1}y^{\frac{d_2}{2}-1}}{2^{\frac{d_1+d_2}{2}}\Gamma\frac{d_1}{2}\Gamma\frac{d_2}{2}}$$

So $(x_1, y_1) \implies (u, v)$ where

$$U = \frac{\frac{x_1}{d_1}}{\frac{x_2}{d_2}}$$

$$v = x_2$$

Then

$$x_1 = \frac{d_1}{d_2} uv$$

$$x_2 = v$$

Now, $x_1, x_2 > 0 \implies u > 0, v > 0$, so we get the Jacobian:\

$$\begin{vmatrix} \frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v} \\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{d_1}{d_2} v & \frac{d_1}{d_2} u \\ 0 & 1 \end{vmatrix} = \frac{d_1}{d_2} v$$

This means that the Joint PDF of u and v is given by:

$$f_{u,v}(u, v) = \frac{e^{-\frac{1}{2}(\frac{d_1}{d_2} uv + v)} (\frac{d_1}{d_2} uv)^{\frac{d_1}{2} - 1} v^{\frac{d_2}{2} - 1} \frac{d_1}{d_2} u}{2^{\frac{d_1 + d_2}{2}} \Gamma \frac{d_1}{2} \Gamma \frac{d_2}{2}}$$

Question 5

We know $f(u, v)$ from 4 where $u > 0$, $v > 0$, so the PDF of u:

$$f_u(u) = \frac{(\frac{d_1}{d_2})^{\frac{d_1}{2}} u^{\frac{d_1}{2} - 1}}{2^{\frac{d_1 + d_2}{2}} \Gamma \frac{d_1}{2} \Gamma \frac{d_2}{2}} \int^{\infty} e^{-\frac{v}{2}(1 + \frac{d_1}{d_2} u) v^{\frac{d_1 + d_2}{2} - 1}} dv$$

Which becomes:

$$= \frac{(\frac{d_1}{d_2}^{\frac{d1}{2}}) u^{\frac{d_1}{2} - 1}}{2^{\frac{d_1 + d_2}{2}} \Gamma \frac{d_1}{2} \Gamma \frac{d_2}{2}} \frac{\Gamma \frac{d_1 + d_2}{2}}{\{\frac{1}{2}(1 + \frac{d_1}{d_2} u)\}^{\frac{d_1 + d_2}{2}}}$$

$$\frac{(\frac{d_1}{d_2})^{\frac{d_1}{2}} u^{\frac{d_1}{2} - 1}}{\beta(\frac{d_1}{2}, \frac{d_2}{2})(1 + \frac{d_1}{d_2} u)^{\frac{d_1 + d_2}{2}}}$$

Where $u > 0$\

This then leads to:

$$f_u(u) = \{ \frac{(\frac{d_1}{d_2})^{\frac{d_1}{2}} u^{\frac{d_1}{2} - 1}}{\beta(\frac{d_1}{2}, \frac{d_2}{2})(1 + \frac{d_1}{d_2} u)^{\frac{d_1 + d_2}{2}}}$$

when $u > 0$ and 0 otherwise. So $U\ F(d_1, d_2)$